J. Symbolic Computation (1988) **5**, 121–129

Thom's Lemma, the Coding of Real Algebraic Numbers and the Computation of the Topology of Semi-algebraic Sets

M. COSTE AND M. F. ROY

IRMAR, Université de Rennes I, Rennes, France

(*Received* 7 *July* 1986)

Thom's lemma, a very simple and basic result in real algebraic geometry, and explained in section 1, has a lot of interesting computational consequences.

We shall outline two of these.

The first one is the fact that a real root ξ of a polynomial P of degree n with real coefficients may be distinguished from the other real roots of P by the signs of the derivatives $P^{(i)}$ of P at ξ, $i = 1, \ldots, n-1$. This offers a new possibility for the coding of real algebraic numbers and for computation with these numbers (see section 2).

The second is based on a generalisation of Thom's lemma to the case of several variables. It gives, after a linear change of coordinates, a cylindric algebraic decomposition of a semi-algebraic set where the incidence relation between the cells is easily obtained (see section 3).

1. Thom's Lemma

DEFINITIONS AND NOTATIONS 1.1: A *sign condition* is any one of the symbols >0, <0 or $=0$. A *generalised sign condition* is any one of the symbols >0, <0, $=0$, $\geqslant 0$, $\leqslant 0$.

If $\varepsilon = (\varepsilon(i))_{i=0,\ldots,n-1}$ is an n-uple of generalised sign conditions, one denotes by $\underline{\varepsilon}$ the n-uple obtained by relaxing the strict inequalities of ε, that is by replacing >0 (resp. <0) by $\geqslant 0$ (resp. $\leqslant 0$).

THOM'S LEMMA

PROPOSITION 1.2. *Let P be a polynomial of degree n with real coefficients, $P', \ldots, P^{(i)}$ its derivatives and $\varepsilon = (\varepsilon(i))_{i=0,\ldots,n-1}$ a n-uple of generalised sign conditions. Let*

$$A(\varepsilon) = \{x \in \mathbb{R} / P^{(i)}(x)\varepsilon(i),\ i = 1, \ldots, n-1\}.$$

Then (a) $A(\varepsilon)$ is either empty or connected, (b) if $A(\varepsilon)$ is non empty the closure of $A(\varepsilon)$ is $A(\underline{\varepsilon})$.

PROOF. Easy, by induction on the degree of P.

REMARK 1.3. (a) In the usual formulation of Thom's lemma (Coste, 1982; Bochnak *et al.*, 1987) ε consists of sign conditions. The form given here is useful for the topological study of a real algebraic curve.

(b) One does not obtain in general the closure of a set defined by strict inequalities by relaxing these inequalities, as is shown by the example of $\{x \in \mathbb{R} | x^3 - x^2 > 0\}$.

0747-7171/88/010121+09 $03.00/0

2. Coding Real Algebraic Numbers

Algorithms proposed up to now to work on real algebraic numbers are semi-numerical: one characterises a real algebraic number ξ by means of a square-free defining polynomial P with integer coefficients and an interval that isolates ξ from all other roots of P. The approach we propose here is purely formal and relies on Thom's lemma.

2.a. CHARACTERISATION OF A REAL ALGEBRAIC NUMBER

PROPOSITION 2.1. *Let P be a polynomial of degree n with integer coefficients. Let ξ and ξ' be two real roots of P. Suppose the signs $\varepsilon(i)$ and $\varepsilon'(i)$ of $P^{(i)}(\xi)$ and $P^{(i)}(\xi')$, $i=1,\ldots,n-1$, are given. Then:*

(i) *If $\varepsilon(i)=\varepsilon'(i)$ for all $i=1,\ldots,n-1$ then $\xi=\xi'$.*

(ii) *In the other case one can decide whether $\xi<\xi'$ or $\xi>\xi'$: let k be the smallest integer such that $\varepsilon(n-k)$ and $\varepsilon'(n-k)$ are different;*

(ii.a) *$\varepsilon(n-k+1)=\varepsilon'(n-k+1)$ is different from $=0$,*

(ii.b) *if $\varepsilon(n-k+1)=\varepsilon'(n-k+1)$ is >0 one has $\xi>\xi'$ if and only if $P^{(n-k)}(\xi)$ is bigger than $P^{(n-k)}(\xi')$ [which is known by looking at $\varepsilon(n-k)$ and $\varepsilon'(n-k)$],*

(ii.c) *if $\varepsilon(n-k+1)=\varepsilon'(n-k+1)$ is <0 one has $\xi<\xi'$ if and only if $P^{(n-k)}(\xi)$ is bigger than $P^{(n-k)}(\xi')$ [which is known by looking at $\varepsilon(n-k)$ and $\varepsilon'(n-k)$].*

PROOF. (i) is a consequence of Thom's lemma. (ii) is also a consequence of Thom's lemma: (ii.a) because of point (i) in Thom's lemma applied to $P^{(n-k+1)}$, (ii.b) and (ii.c) because the set $\{x\in\mathbb{R}/P^{(i)}(x)\varepsilon(i),\ i=n-k+1,\ldots,n-1\}$ is connected by Thom's lemma applied to $P^{(n-k+1)}$, and that on a connected set, the sign of the derivative of a polynomial gives its behaviour (increasing or decreasing).

2.b. ALGORITHMS FOR COMPUTING ON REAL ALGEBRAIC NUMBERS

In section 2b let P be a polynomial of degree n with integer coefficients having only simple roots and d real roots. It is clear that Proposition 2.1 gives a way of coding a real algebraic number root of P (by the sign it gives to the derivatives of P) and that, given such a code, the sign of any polynomial Q with integer coefficients at a given real algebraic number is fixed. We propose in the following an algorithm for computing the codes of the roots of P and more generally algorithms for computing sign conditions given by roots of P to other polynomials.

All algorithms are based on a generalisation of Sturm theorem due to Sylvester (1853) (algorithm b_1), revisited by Ben-Or, Kozen & Reif (1986) (algorithm b_3).

Let us introduce some notation. Let $R, Q_1,\ldots,Q_k$ be polynomials with integer coefficients and let $\varepsilon=(\varepsilon(1),\ldots,\varepsilon(k))$ be a sequence of sign conditions $(>0, <0, =0)$. One denotes by $c_\varepsilon(R; Q_1,\ldots,Q_k)$ the number of real roots of R giving to $Q_1,\ldots,Q_k$ the signs $\varepsilon(1),\ldots,\varepsilon(k)$.

Let P_0 be R, P_1 be Q, P_{i+1} be the opposite of the remainder of the division of P_{i-1} by P_i. One calls the sequence of polynomials so obtained the *generalised Sturm sequence associated with R and Q* and one denotes by $v_{R,Q}(-\infty)$ and $v_{R,Q}(+\infty)$ the number of sign changes in the sequences of the signs of the P_i at $-\infty$ and $+\infty$. The *Sturm sequence of P* is the generalised Sturm sequence associated to P and P'.

PROPOSITION 2.2. *With the above definitions and notations, if R and Q are coprime, one has* $c_{>0}(R; Q) - c_{<0}(R; Q) = v_{R, P'Q}(-\infty) - v_{R, P'Q}(+\infty)$.

PROOF. Similar to the proof of Sturm theorem (see [S]).

2.b.1. *Generalised Sturm algorithm for an inequality (algorithm b_1)*

Let Q be a polynomial with integer coefficients, prime to P. The aim of the algorithm is to compute $c_{>0}(P; Q)$ and $c_{<0}(P; Q)$. Let $v(1) = v_{P, P'}(-\infty) - v_{P, P'}(+\infty)$ and $v(2) = v_{P, P'Q}(-\infty) - v_{P, P'Q}(+\infty)$. Using Proposition 2.2 we have

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_{>0}(P; Q) \\ c_{<0}(P; Q) \end{bmatrix} = \begin{bmatrix} v(1) \\ v(2) \end{bmatrix}$$

We can easily deduce $c_{>0}(P; Q)$ and $c_{<0}(P; Q)$ from $v(1)$ and $v(2)$.

2.b.2. *Algorithm for computing the number of real roots of P giving a fixed sign condition (>0, <0, $=0$) to a polynomial Q (algorithm b_2)*

Let Q be a polynomial with integer coefficients. The aim of the algorithm is to compute $c_{>0}(P; Q)$, $c_{<0}(P; Q)$ and $c_{=0}(P; Q)$.

One computes the GCD R of P and Q. If R is of degree 0 one applies b_1. If R is of strictly positive degree, one defines $S = P/R$. The number $c_{=0}(P; Q)$ is equal to the number of real roots of R (computed by Sturm), $c_{>0}(P; Q)$ is equal to $c_{>0}(S; Q)$ and $c_{<0}(P; Q)$ is equal to $c_{<0}(S; Q)$. One computes then $c_{>0}(S; Q)$ and $c_{<0}(S; Q)$ b_1.

2.b.3. *Generalised Sturm algorithm for several inequalities (algorithm b_3)*

Let $Q_1, \ldots, Q_k$ be polynomials in one variable with integer coefficients. The *sign conditions realised by $Q_1, \ldots, Q_k$ at the real roots of P* are the k-uples of signs conditions ε such that $\{x \in \mathbb{R} | P(x) = 0$ and $Q_j(x) \in (j), j = 1, \ldots, k\}$ is *non empty*.

Let us suppose now that the Q_j are prime to P. The aim of the algorithm is to determine the number of real roots of P giving to $Q_1, \ldots, Q_k$ fixed signs. The idea is to compute the generalised Sturm sequences associated to the products of the subsets of $\{Q_1, \ldots, Q_k\}$ and use the sign variations of these generalised Sturm sequences and some linear algebra (generalising b_1) to deduce the values of the $c_\varepsilon(P; Q_1, \ldots, Q_k)$. If done without care, this leads to an exponential algorithm, since we have 2^k subsets of $\{Q_1, \ldots, Q_k\}$, hence 2^k generalised Sturm sequences to compute, to get the values c_ε for the 2^k-sign conditions ε. To avoid this exponential growth the idea (Ben-Or *et al.*, 1986) is to remark that the number $d(k)$ of distinct sign conditions realised by $Q_1, \ldots, Q_k$ at the real roots of P is, for any k, smaller than the number d of real roots of P.

More precisely one wants to determine the sign conditions $\varepsilon_{k,1}, \ldots, \varepsilon_{k,d(k)}$, with $d(k) \leqslant d$, realised by $Q_1, \ldots, Q_k$ at the real roots of P and the number $c(k, j) = c_{\varepsilon_{k,j}}(P; Q_1, \ldots, Q_k)$ of elements of the non empty set

$$\{x \in \mathbb{R} | P(x) = 0 \text{ and } Q_j(x) \varepsilon_{k,j}(i),\ i = 1, \ldots, k\}.$$

We are going to compute by induction on k

(1_k) the list $\varepsilon_k = (\varepsilon_{k,1}, \ldots, \varepsilon_{k,d(k)})$, $d(k) \leqslant d$, of the k-uples of sign conditions realised by $Q_1, \ldots, Q_k$ at the real roots of P,

(2_k) the numbers $c(k, 1), \ldots, c(k, d(k))$

(3_k) a list of $S_k = (S_{k,1}, \ldots, S_{k,d(k)})$ of $d(k)$ subsets of $\{1, \ldots, k\}$,

(4_k) for each $j = 1, \ldots, k$ the number $v(k,j) = v_{P,Q_{k,j}}(-\infty) - v_{P,Q_{k,j}}(+\infty)$, with $Q_{k,j}$ the product of the Q_l for $l \in S_{k,j}$, and the list of the $Q_{k,j}$,

(5_k) an invertible matrix $\mathbf{A}_k$ of dimension $d(k)$ such that $\mathbf{A}_k . c = v$, where c is the vector $(c(k,1), \ldots, c(k,d(k)))$ and v the vector $(v(k,1), \ldots, v(k,d(k)))$.

The case $k = 0$ is given by $d(0) = 0$.

Suppose that we are given $(1_k), \ldots, (5_k)$ and consider the case $k+1$.

We compute $d(k)$ generalised Sturm sequences associated to the $Q_{k,d(k)+j} = Q_{k+1}Q_{k,j}$, $j = 1, \ldots, d(k)$ and $v(k,l) = v_{P^{\prime}, Q_{k,1}}(-\infty) - v_{P^{\prime}, Q_{k,1}}(+\infty)$, $l = d(k)+1, \ldots, 2d(k)$. One considers the $2d(k)$ dimension matrix

$$\mathbf{A}' = \begin{bmatrix} A_k & A_k \\ A_k & -A_k \end{bmatrix},$$

and the equality $\mathbf{A}' . c' = v'$, obtained from (5_k) and 2.2 where c' is the vector

$$(c_{\varepsilon_{k,j}>0}(P; Q_1, \ldots, Q_{k+1}),\ j = 1, \ldots, d(k), \quad c_{\varepsilon_{k,j}<0}(P; Q_1, \ldots, Q_{k+1}),\ j = 1, \ldots, d(k))$$

and V' the vector $(v(k,1), \ldots, v(k,2d(k)))$.

The matrix $\mathbf{A}'$ is invertible. One computes c' by inverting $\mathbf{A}'$.

Let us define $d(k+1)$ as the number ($\leqslant d$) of non zero elements in c' and let

$$m_1, \ldots, m_{d(k+1)} (m_1 < \ldots < m_{d(k+1)} \leqslant 2d(k))$$

be the numbers j such that the jth coordinate of c' is different from 0. The sign conditions $\varepsilon_{k+1,1}, \ldots, \varepsilon_{k+1,d(k+1)}$ realised by $Q_1, \ldots, Q_{k+1}$ at the real roots of P are the m_jth elements, $j = 1, \ldots, d(k+1)$, of the list

$$((\varepsilon_{k,1}, >0), \ldots, (\varepsilon_{k,d(k)}, >0), (\varepsilon_{k,1}, <0), \ldots, (\varepsilon_{k,d(k)}, <0)).$$

One can then extract from $\mathbf{A}'$ the columns of number $m_1, \ldots, m_{d(k+1)}$, which gives a $2d(k), d(k+1)$ matrix $\mathbf{A}''$. An invertible matrix $\mathbf{A}_{k+1}$, of dimension $d(k+1)$ can be extracted from $\mathbf{A}''$. Let us denote $l_1, \ldots, l_{d(k+1)}$, $l_1 < \ldots < l_{d(k+1)}$, the elements of $\{1, \ldots, 2d(k+1)\}$ such that the l_ith line of $\mathbf{A}''$ appears in $\mathbf{A}_{k+1}$.

The list $(S_{k+1,1}, \ldots, S_{k+1,d(k+1)})$ is obtained taking in the list

$$(S_{k,1}, \ldots, S_{k,d(k)}, S_{k,1} \cup \{k+1\}, \ldots, S_{k,d(k)} \cup \{k+1\})$$

the elements of number $l_1, \ldots, l_{d(k+1)}$.

It is easy to verify $(1_{k+1}), \ldots, (5_{k+1})$.

2.b.4. *Algorithm for determining the number of real roots of P giving fixed signs to polynomials $Q_1, \ldots, Q_k$ (algorithm b_4)*

Let $Q_1, \ldots, Q_k$ be polynomials with integer coefficients. The aim of the algorithm is to determine the number of real roots of P giving to $Q_1, \ldots, Q_k$ fixed signs. More precisely, we want to determine

($1'_k$) k-uples of signs $\varepsilon_{k,1}, \ldots, \varepsilon_{k,m}$ (with $m \leqslant d$) such that the $c_{\varepsilon_{k,j}}(P; Q_1, \ldots, Q_k)$ are different from zero;

($2'_k$) for each $\varepsilon_{k,j}$ (if one denotes $Q_{j,1}, \ldots, Q_{j,k'}$ the k' polynomials among $(Q_i)_{i=1,\ldots,k}$ such that $\varepsilon_{k,j}(i)$ is not zero and $\varepsilon'_{k,j}$ the k'-uple of strict signs on the $Q_{j,l}$ taken from $\varepsilon_{k,j}$), a polynomial S_j dividing P, prime with all the $Q_{j,l}$s and such that $c_{\varepsilon_{k,j}}(S_j; Q_{1,j}, \ldots, Q_{k,j})$ is equal to $c_{\varepsilon_{k,j}}(P; Q_1, \ldots, Q_k)$.

The case $k=1$ *is* b_2.

Let us suppose the problem solved for $P, Q_1, \ldots, Q_k$ and let us add the polynomial Q_{k+1}. One considers for each k-uple of signs $(\varepsilon_{k,j})_{j=1,\ldots,m}$ given by (1_k) for $P, Q_1, \ldots, Q_k$ the GCD R of Q_{k+1} and S_j. One applies b_3 to $S_j/R, Q_{j,1}, \ldots, Q_{j,k'}, Q_{k+1}$, which gives the result for strict signs $(>0, <0)$ on Q_{k+1}. For the case where the sign of Q_{k+1} is 0, one applies again b_3 to $R, Q_{j,1}, \ldots, Q_{j,k'}$.

2.b.5. *Algorithm for characterising the real roots of a polynomial (algorithm* b_5)

One wants to compute the signs given to the derivatives of P by the real roots of P. One applies b_4 to $P, P', \ldots, P^{(n-1)}$.

As seen in Proposition 2.1 a real root ξ of P can be coded by the sequence of signs it gives to the derivatives of P.

2.b.6. *Algorithm for deciding the sign of a polynomial with integer coefficients at a real algebraic number*

One supposes the real algebraic number ξ, root of P, coded as in b_5 and one wants to compute the sign of Q in ξ.

One applies b_4 to $P, P', \ldots, P^{(n-1)}, Q$.

2.c. COMPLEXITY OF THE ALGORITHMS

The precise complexity study of the algorithm is not included in the present paper and is in preparation (Roy and Szpirglas, in prep.).

Let us indicate quickly why algorithm b_3 is a polynomial-time algorithm (then clearly $b_4, \ldots, b_7$ will also be polynomial-time algorithms). If R is a polynomial in one variable with integer coefficients, let $C(R)$ be the maximum of the degree of R and of the length of the coefficients of R, written in base 2. Let $P, Q_1, \ldots, Q_k$, be as in b_3, let $N = \sup(k, C(P), C(Q_i),\ i=1, \ldots, k)$, let (step i) be the passage from $Q_1, \ldots, Q_i$ to $Q_1, \ldots, Q_{i+1}$ in b_3. The number of generalised Sturm sequences to compute in (step i) is bounded by d, where d is the number of real roots of P, hence by N. All the computations in (step i) are polynomial in N if one uses the theory of subresultants (for example, see Loos, 1982) for the computation of generalised Sturm sequences.

2.d. GENERALISATION TO SEVERAL VARIABLES

It is not always the case that a real algebraic number is given as a root of a polynomial with integer coefficients. In many natural geometric situations it happens that it is given as a root of a polynomial having itself real algebraic coefficients. Of course, it is in principle possible, using the theorem of the primitive element, to compute a polynomial with integer coefficients having the real algebraic number as a root. It we want to avoid these primitive element computations, we can proceed as follows.

The coding of a real algebraic number given as a root of a polynomial P with real algebraic coefficients by the signs of the derivatives of P will be a particular case of the following general decision problem (*): let us consider

$$P_1 \in \mathbb{Q}[T_1],\ P_2 \in \mathbb{Q}[T_1, T_2], \ldots, P_n \in \mathbb{Q}[T_1, \ldots, T_n],\ Q_1, \ldots, Q_k \in \mathbb{Q}[T_1, \ldots, T_n];$$

compute the sign conditions $\varepsilon = (\varepsilon(1), \ldots, \varepsilon(k))$ such that there exists a real root ξ_1 of P_1,

there exists a real root ξ_2 of $P_2[\xi_1, T_2], \ldots$, there exists a real root ξ_n of $P[\xi_1, \xi_2, \ldots, \xi_{n-1}, T_n]$ such that $Q_i(\xi_1, \xi_2, \ldots, \xi_{n-1}, \xi_n)\varepsilon(i)$, and for each such ε the number of roots ξ_1 of P_1 such that there exists a fixed number m_2 of roots of $P(\xi_1, T_2) \ldots$ such that there exists a fixed number of roots m_n of $P[\xi_1, \xi_2, \ldots, \xi_{n-1}, T_n]$ such that $Q_i(\xi_1, \xi_2, \ldots, \xi_{n-1}, \xi_n)\varepsilon(i)$ for all $i = 1, \ldots, k$.

The case $n = 1$ corresponds obviously to b_4. The case n is treated recursively from case $n-1$, by considering the polynomials $P_n, Q_1, \ldots, Q_k$ as polynomials in the variable T_n. Computing the generalised Sturm sequences corresponding to the situation, we get, by considering signs conditions on leading terms on generalised Sturm sequences, a new problem (*) in the variables $T_1, \ldots, T_{n-1}$.

2.e REMARK

The algorithms proposed can be used in any computable real closed fields, for example in the field of real Puiseux series which are algebraic over $\mathbb{Q}(X)$, which appears in natural geometric situations. Since these fields are not archimedean, the technique of dichotomy for the isolation of real roots can no more be applied here.

2.f. APPLICATION TO THE COMPUTATION OF THE TOPOLOGY OF A CURVE

This subject will be developed in another paper (Roy, 1987). Let us simply say that the computation of the topology of a curve given by the equation $Q(x, y) = 0$ (number of connected components, isotopy type of the embedding of the curve in the real plane, number of multiple points and of branches passing through them, number of real isolated points) can be made if one knows the answer to the following questions

(1) determine above the real roots ξ of the discriminant D of $Q(X, Y)$ with respect to Y the number of real roots of $Q(\xi, Y)$,
(2) determine just to the left and to the right of the real roots of the discriminant D the number of real half-branches and how they glue over ξ.

Using generalised Sturm theorem and Thom's lemma one can show that answering these questions is equivalent to characterising the real roots of D (in the sense of b_5) and, using 2.d, evaluating the signs of other polynomials at these real roots (by b_6, which is a special case of b_3).

3. Computing the Topology of Semi-algebraic Sets

3.a. CYLINDRIC ALGEBRAIC DECOMPOSITION

We recall that a semi-algebraic subset of $\mathbb{R}^n$ is a subset defined by a boolean combination of equations and inequalities involving a finite number of polynomials in $\mathbb{R}[X_1, \ldots, X_n]$. Any systematic study of semi-algebraic sets is based on the following result (Theorem 3.2). We introduce first some terminology from Schwartz & Sharir (1983).

DEFINITION 3.1. Let $T \subset \mathbb{R}[X_1, \ldots, X_n]$ be a family of polynomials and $A \subset \mathbb{R}^k$ any subset. We say that A is *T-invariant* when any polynomial $f \in T$ is either strictly positive, or strictly negative, or identically zero on the whole of A.

THEOREM 3.2. *For any finite family* $S \subset \mathbb{R}[X_1, \ldots, X_n]$ *one can produce a finite family* $T \subset \mathbb{R}[X_1, \ldots, X_{n-1}]$ *such that for any connected subset* $A \subset \mathbb{R}^{n-1}$ *which is T-invariant, there are an integer* $q \in \mathbb{N}$ *and continuous functions*

$$\xi_1 < \ldots < \xi_q : A \to \mathbb{R}$$

which give, for any $a \in A$, *all the real roots of all the polynomials* $f(a, X_n), f \in S$. *Moreover, if A is semi-algebraic, the graphs of the* ξ_i*s and the slices of the cylinder* $A \times \mathbb{R}$ *cut out by these graphs are semi-algebraic.*

PROOF. See, for example, Bochnak *et al.* (1987), ch. 2, th. 2.3.1.

DEFINITION 3.3. In the situation of the preceding theorem, we say that the family T *cylindrifies* the family S.

Theorem 3.2 above goes back to Lojasiewicz (1965, pp. 105–110), which contains the first systematic study of semi-algebraic sets. Now, given a finite family $S \subset \mathbb{R}[X_1, \ldots, X_n]$, this result yields by induction on n a finite decomposition of $\mathbb{R}^n$ into semi-algebraic cells C_i, $i = 1, \ldots, p$, such that each C_i is homeomorphic to $]0, 1[^{d_i}$ for some d_i, and is S-invariant. Collins has given the name of *cylindric algebraic decomposition* (c.a.d.) to such a decomposition of $\mathbb{R}^n$, and he has described an algorithm for the production of c.a.d. (cf. Collins, 1975).

3.b. INCIDENCE RELATION AND STRATIFYING FAMILIES

Schwartz & Sharir (1983) have shown that, up to a linear change of coordinates, the c.a.d. has the property that the closure of any cell is a union of cells. Then an algorithm for determining the incidence relation between the cells would give the possibility to calculate all the topological information needed on semi-algebraic sets. For instance, the number and the description of connected components, a triangulation, the homology groups, etc. Schwartz and Sharir propose an algorithm for the incidence relation, but this algorithm is rather complicated except in the case when the dimension of the cells are n and $n-1$, respectively, which is all they need for their purpose.

Here we want to sketch how a result known to real algebraic geometers and which is a generalisation of Thom's lemma to the case of several variables provides an alternative algorithm for the incidence relation. This generalisation is essentially due to Efroymson (1976). A proof may be found in Coste (1982, th. 3.2), and a more detailed version is in Bocknak *et al.* (1987, ch. 9, sect. 1). We start with a definition.

DEFINITION 3.4. A *stratifying family* of polynomials in $\mathbb{R}[X_1, \ldots, X_n]$ is a family $(f_{i,j})$, $i = 1, \ldots, n, j = 1, \ldots, l_i$, such that:

(i) For any (i, j) the polynomial $f_{i,j}$ is in $\mathbb{R}[X_1, \ldots, X_i]$ and its leading coefficient (as a polynomial in X_i) is a non zero constant.
(ii) For any fixed i, the family $(f_{i,j})$, $j = 1, \ldots, l_i$, is stable under derivation with respect to X_i (that is, the derivative of $f_{i,j}$ with respect to X_i is either zero or one of the $f_{i,j'}$, $j' \neq j$).
(iii) For any $i < n$, the family $(f_{i,j})$, $j = 1, \ldots, l_i$, cylindrifies the family $(f_{i+1,j})$, $j = 1, \ldots, l_{i+1}$.

A stratifying family yields a c.a.d. which is very satisfying.

THEOREM 3.5 (GENERALISED THOM'S LEMMA). *Let $(f_{i,j})$, $i=1,\ldots,n$, $j=1,\ldots,l(i)$, be a stratifying family of polynomials. Consider the finite decomposition of $\mathbb{R}^n$ given by all the non-empty semi-algebraic sets of the form*

$$C_\varepsilon = \bigcap_{i=1}^{p} \bigcap_{j=1}^{l(i)} \{x \in \mathbb{R}^n | \text{sign}\,(f_{i,j}(x)) \varepsilon(i,j)\}$$

where $\varepsilon = (\varepsilon(i,j))$ is a family of sign conditions <0, $=0$ or >0.

(i) Any such C is homeomorphic to $]0,1[^d$ for some d.

(ii) The closure of any such C is obtained by relaxing the strict inequalities which appear in its definition.

PROOF. See Bocknak *et al.* (1987, ch. 9, th. 9.1.4).

So for the c.a.d. associated to a stratifying family the formulas describing the cells and the incidence relation between the cells are given free. A cell C is contained in the closure of a cell C' if and only if any sign condition on the $f_{i,j}$s which appears in the description of C also appears in the description of C'. The only problem left is to determine which collections of signs $(\varepsilon(i,j))$ give non empty cells. One has to use here a going-up process by induction on the dimension for which the algorithms for computing with real algebraic numbers sketched in 2 can be useful.

Finally, after a linear change of coordinates, any semi-algebraic set can be decomposed via a c.a.d. given by a stratifying family.

PROPOSITION 3.6. *Let $g_1,\ldots,g_k$ be polynomials in $\mathbb{R}[X_1,\ldots,X_n]$. Then one can produce a linear change of coordinates $u\colon \mathbb{R}^n \to \mathbb{R}^n$ and a stratifying family of polynomials $(f_{i,j})$ in $\mathbb{R}[X_1,\ldots,X_n]$ such that $g_j(u(X_1,\ldots,X_n)) = f_{n,j}(X_1,\ldots,X_n)$ for $j=1,\ldots,l_n$.*

PROOF. The properties (i) and (ii) in the definition of the stratifying families (3.4) are easily obtained (Bocknak *et al.*, ch. 9, prop. 9.1.2). For the property (iii) of 3.4 one may use the algorithm of "augmented projection" in Collins (1975).

REMARK 3.7. The introduction of derivatives multiplies the number of cells of the c.a.d. But these cells may be grouped together, in the going-up process of the induction on the dimension, in semi-algebraic connected subsets of $\mathbb{R}^i$ over which the polynomials $(f_{i+1,j})$, $j=1,\ldots,l_{i+1}$, have a constant number of roots, since the incidence relation in the c.a.d. of $\mathbb{R}^i$ is known. This remark is due to D. Lazard. We do not know if this remark improves the complexity of the cylindrical algebraic decomposition.

Let us remark also that, in some situations, when we want to get a description of the projection of a semi-algebraic set in a given direction, a linear change of coordinates is forbidden and the stratifying families cannot be used.

References

Ben-Or, M., Kozen, D., Reif, J. (1986). The complexity of elementary algebra and geometry. *J. Computation Systems Sci.* **32**, 251–264.

Bochnak, J., Coste, M., Roy, M.-F. (1987). *Geométrie algébrique réelle. Ergebnisse der Mathematik*, Springer-Verlag, Berlin.

Collins, G. (1975). Quantifier elimination for real closed fields by cylindric algebraic decomposition. In: *Second GI Conference on Automata Theory and Formal Languages. Lecture Notes in Computer Science, 33.* Springer-Verlag, Berlin, pp. 134–183.

Coste, M. (1982). Ensembles semi-algébrique. In: *Géométrie algébrique réelle et formes quadratiques. Lecture Notes in Mathematics, 959.* Springer-Verlag, Berlin, pp. 109–138.
Efroymson, G. (1976). Substitution in Nash functions. *Pac. J. Math.* **54,** 109–138.
Lojasiewicz, S. (1965). Ensembles semi-analytiques. *Lecture Notes, I.H.E.S.*
Loos, R. (1982). Generalized polynomial remainder sequences. In: *Computer Algebra, Symbolic and Algebraic Computation.* Springer-Verlag, Berlin.
Roy, M.-F. (1987) Computation of the topology of a real algebraic curve. Congress on "Computational geometry and topology and computation in teaching mathematics", Seville (to appear).
Roy, M.-F., Szpirglas, A. (In prep.). Complexity of computation on real algebraic numbers.
Schwartz, J. T., Sharir, M. (1983). On the "Piano mover's" problem. II. *Advan. Applied Math.* **4,** 298–351.
Sylvester, J. J. (1853). On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's function. *Trans. Roy. Soc. London.*